DOCUMENT RESUME

ED 389 229 FL 023 449

AUTHOR Orr, Thomas; And Others
TITLE An Analysis of Lexical Frequency and Discourse Need for Instructive Purposes. Technical Report 95-5-002.
INSTITUTION Aizu Univ., Aizuwakamatsu (Japan). Center for Language Research.
PUB DATE 17 Apr 95
NOTE 16p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Computer Science; Computer Software; Discourse Analysis; *Educational Needs; English (Second Language); *English for Special Purposes; Foreign Countries; *Freshman Composition; Higher Education; Second Language Instruction; Technical Writing; *Vocabulary; *Word Frequency; Writing Instruction

ABSTRACT

The report details a study to identify vocabulary needed by Japanese students of English as a Second Language to function successfully in college computer science courses and research laboratory apprenticeships. The vocabulary was then to be taught in the first two semesters of freshman English composition. The study involved development of simple computer software and administration of an electronic questionnaire and individual interviews. The software was designed to study lexical frequency in the published annual review of the university's school of computer science and engineering. Software specifications are given here. The resulting vocabulary is also presented, in four categories: university academic vocabulary; university physical environment vocabulary; people-related vocabulary; and computer science discourse vocabulary. It was found that most of the words students need to know to negotiate computer science content materials in English are words common to general English rather than words of a semi- of highly technical nature. Contains 17 references. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 389 229

Technical Report 95-5-002

An Analysis of Lexical Frequency and Discourse Need for Instructive Purposes

Thomas Orr, Kiel Christianson,
Christian Goetze, Hideaki Okawara

April 17, 1995

BEST COPY AVAILABLE

Center for Language Research
University of Aizu
Aizuwakamatsu, Fukushima
965-80 Japan

FL023449

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Thomas Orr
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

The technical reports are published for early dissemination of research results by the members of the University of Aizu. The completed results may be submitted later to journals and conferences for publication.

Title:

An analysis of lexical frequency and discourse need for instructive purposes

Authors:

Thomas Orr, Kiel Christianson, Christian Goetze, Hideaki Okawara

Key Words and Phrases:

English for Specific Purposes (ESP), English for Science and Technology (EST), lexical frequency, academic vocabulary, computer science vocabulary, needs analysis

Abstract:

The research activities detailed in this technical report consist of simple research software development and application along with electronic questionnaires and face-to-face interviews for the purpose of identifying a corpus of English words to be taught in freshman Composition 1 and 2. It is believed that mastery of these words is necessary for University of Aizu students to be able to function successfully in their computer courses and research lab apprenticeships.

Report Date:

April 17, 1995

Written Language:

English

Any Other Identifying Information of this Report:

Distribution Statement:

First Issue: 100 copies

Supplementary Notes:

This research was conducted by the Composition Courseware Research Team as part of its responsibilities to develop effective composition instruction for University of Aizu freshmen.

Center for Language Research
University of Aizu
Aizu-Wakamatsu City
Fukushima 965-80 Japan

An Analysis of Lexical Frequency and Discourse Need for Instructive Purposes

Thomas Orr, Kiel Christianson,
Christian Goetze, Hideaki Okawara

1 Introduction

The defining characteristic of an ESP[1] approach to English language instruction is its thorough and continual needs analysis which identifies target language needs and facilitates the formation of instructional priorities. The Center for Language Research (CLR) at the University of Aizu exemplifies this trend in language education as it studies the English language needs of students and professionals in computer science and develops instruction material and services to enable nonnative speakers of English (NNSs) to function more successfully within the international computer science discipline.

One of the CLR teams active in needs analysis is the English Composition Courseware Team,[2] which is responsible for developing the syllabus for the first two semesters of student writing instruction. The team's research activity detailed in this technical report consists of simple research software development and application along with electronic questionnaires and face-to-face interviews for the purpose of identifying a prospective English corpus for instructive purposes. These words would be taught in Composition 1 and 2 to assist University of Aizu students function successfully in their computer courses and research lab apprenticeships.

Though knowledge of the English vocabulary listed in this report will be essential for freshman students at this university, the authors recognize that the material covered here may also serve broader needs. First, the English vocabulary will be extremely useful to high school students preparing for coursework at the University of Aizu; and second, the discussion of research methodology will be useful to ESP educators who wish to initiate vocabulary research at their own locale. It is, thus, with a desire to provide genuinely useful material to both students and educators that we offer this technical report to our readers.

[1] English for Specific Purposes

[2] The members of the English Composition Courseware Team are Associate Professor Thomas Orr (Center for Language Research), Assistant Professor Kiel Christianson (Center for Language Research), Assistant Professor Christian Goetze (Department of Computer Software), and special research assistant Hideaki Okawara (University of Aizu student).

2 Research Method

2.1 Background

Lexical needs analysis is not new to ESP; however, it is unfortunately far too rare in most language programs. In contrast to English for General Purposes, English for Specific Purposes and its major subdomains (English for Medical Purposes, English for Business Purposes, English for Academic Purposes, English for Vocational Purposes, and English for Science and Technology) seek to orient language learners to a specific subculture characterized by specific language used to accomplish specific tasks. As such, the identification of an English corpus used to convey a field's content and used to accomplish its goals is both reasonable and necessary if a language program is to be effective.

Some of the early work in corpus construction and analysis begun in the 1960s (Kucera & Francis, 1967) has been carried further by ESP specialists in several academic fields. Salager-Meyer (1983) and Chandler-Burns (1986, 1987, 1994a, 1994b, 1995), for example, have been actively applying computational linguistics in the field of medicine; Inman (1978) conducted a well-known lexical analysis of scientific and technical prose; and Bramki & Williams (1984) have studied the vocabulary of economics. In addition to studies of discipline-specific lexicons, some language programs have carried out their own vocabulary studies of language within the local university environment in order to prioritize their own English language instruction. Flowerdew's 1993 study at the Sultan Qaboos University is one good example. These kinds of studies have been important contributions to ESP and must continue if programs are to effectively prioritize their instruction so that students can learn to perform the most needful language tasks within the time alloted to their language studies.

2.2 Knowledge Sought

The goal of this research project was to identify the English words most needed by our students for them to function well in the academic setting here at the University of Aizu. These words would be organized under four categories: 1) the vocabulary of academia (to understand classroom or research laboratory instruction); 2) the vocabulary of campus locations and objects (to understand references to the physical university environment); 3) the vocabulary of university people (to understand references to local groups on campus), and 4) the vocabulary of introductory computer science discourse (to understand written and spoken communication about computer science). These four lists, however, required four different research tasks.

The first list (Academic Vocabulary) required a survey of University faculty to identify which words were felt to be most critical for students to understand instruction in lectures and joint research projects with the faculty. These items were gathered via an electronic survey on the campus network and through discussions over a period of two years with faculty who had experience teaching our students.

The second list (University of Aizu Physical Environment Vocabulary) required the consultation of campus maps and floor plans, university email announcements, and a simple analysis of student writing during the first two years of freshman Composition 1 and 2 to learn what objects in the physical environment were most often referred to and which items gave local NNSs the most difficulty.

The third list (University of Aizu People Vocabulary) required the consultation of the *University of Aizu Regulations* and the application of native-speaker knowledge of standard English

usage in American universities.

The fourth list (University of Aizu Computer Science Discourse Vocabulary) required the most thought and technical assistance. The first problem was to identify what language records would best represent the vocabulary necessary to understand the verbal and written discourse of the local computer science culture, and the second problem was how to extract the appropriate lexicon most efficiently. To solve these two problems, the research team decided that the development of software to perform a simple computational analysis of lexical frequency on the *University of Aizu 1993 Annual Review: School of Computer Science and Engineering* would be the wisest choice for the following reasons: 1) the Annual Review contained the best English textual record of descriptions of the university, the departments, the centers, and the research labs; 2) it contained abstracts of all faculty research work during the 1993 academic year; and 3) the document was on the campus network, which made electronic investigation very easy.

Though the team had originally considered whether or not to include course textbooks, campus email discussions, seminar announcements & abstracts, and university technical reports, it was decided that the Annual Review would be the one document that gave the most complete and most accurate record of local computer science discourse that would include both profiles of university programs and discussions of specific university research activities. Once the target text was agreed upon, the next step was to design a software tool for the network that could easily abstract the needed data.

2.3 Software Sought

The software tool for a lexical frequency study needed to be able to do the following at the touch of one button:

- give a total count of words within the text
- list all words in order of frequency followed by the number of times the word appeared in the text, e.g., *word (number)*
- list all words in alphabetical order followed by the number of times the word appeared in the text, e.g., *A-word (number), B-word (number)*
- record this data at the end of the text so it could be printed along with the text if desired

Professor Goetze, the team's technical assistant from the Department of Computer Software responded to this request with the following software design.

2.3.1 Software Specifications

Even though the task of counting words and sentences may appear relatively easy, design software that can accomplish this in a sufficiently precise manner can be amazingly difficult. The main problem stems from distinguishing between abbreviations and words at the end of sentences. Minor problems are encountered when one processes "semi-formatted" text with itemized or numbered lists embedded in the text.

In this project, the following definitions were used:

1. A *white space* is either a space, a line break or the beginning or the end of the text.

2. A *word* is a sequence of alphabetical characters that may contain the special characters "-" and "'", and that must contain at least one vowel.

3. A *sentence* is a sequence of *words* terminated by ".", "...", "!" or "?". There must be a *word* immediately preceding those characters, and those characters must be followed by a *white space*.

4. Any sequence of non-*white space* characters following any one of "/", ".", "\", "#", "@" are ignored.

Definition 2 is a heuristic to distinguish between abbreviations and "real" words. This definition has a fairly good hit-ratio, and is very "cheap" to implement. Using dictionaries would work better, but is more "expensive" to implement.

Definition 3 is also mainly geared at distinguishing between abbreviations and words. One wants to ignore embedded dots as those are often used in e-mail addresses and other words in computer-related literature.

Definition 4 filters out most formatting commands used by different text processing systems (LaTeX, TeX, nroff).

The implementation was done in C, and comprises about 200 lines of code. The program works as a filter, taking the text to be examined as input and producing a statistic as output. Since a hash-table is used for the word-frequency test, the run-time performance is basically proportional to the size of the text, where text with rich vocabulary might take slightly longer than the same size text with a simple repertoire.

The program can be used as is or via text editors like "vi" or "emacs". The program and editor-scripts can be downloaded via:

ftp://ftp.u-aizu.ac.jp/u-aizu/clr/count-words.tar.gz

The following shows a sample of the data from the *Annual Review* after the software was employed:

```
    vocabulary = 7683
         words = 51983

the(2724), of(2481), and(1778), a(1203), in(1083), to(771), is(715), for(641),
on(525), with(328), are(323), we(264), by(259), this(256), as(253), an(238),
computer(237), that(225), research(185), be(181), systems(181), system(171),
laboratory(164), software(156), university(150), design(147), which(146),
model(143), can(135), professor(125), parallel(124), new(123), method(121),
has(120), based(119), processing(113), from(111), achievement(110), have(110),
it(109), summary(109), also(103), using(101), data(97), international(97),
algorithm(96),...

a(1203), ability(5), able(5), abnormal(1), about(18), above(7), abroad(1),
abrupt(1), absence(1), absorption(6), abstract(15), abstracted(1),
abstracting(1),abstraction(3), abstractly(1),...
```

3 Results

The following sections list the corpora that resulted from this study and are thought to be of value to University of Aizu students preparing for English instruction and computer science discourse during their freshman year in Composition 1 and 2. The lists do not contain ALL the words that students need to learn during their four-year course of study, but only those words that would be appropriate for mastery during their freshman year of studies.

3.1 University of Aizu Academic Vocabulary

absence(s); (to be) absent (from)
abstract (of an article)
analysis (analyses); to analyze; analytic, analytical; analytically
answer(s); to answer (a question)
(to make) an appointment (appointments)
(to give) an assignment (assignments); to assign
(to take) attendance; to attend (class, a conference, a meeting, a seminar)
to call on (someone)
class(es) (= a group of students or a course)
(academic) conference(s)
(to make) a correction (corrections); to correct; correct (answers); (to do something) correctly
course(s)
coursework
(to set, to meet) a deadline (deadlines)
(to make) a deletion (deletions); to delete
(to give) direction(s); to direct; (a) direct (person); (to speak) directly
(to have) a discussion (discussions); to discuss
(to be) due
(to get an) education; to educate
to erase
(to make) an error (errors)
evaluation(s); to evaluate
(to give, to have, to take) an exam(ination) (exam(ination)s); to examine
(to give) an example (examples); to exemplify; exemplary
grade(s); to grade
to hand in (something), to hand (something) in
(to distribute) a handout (handouts); to hand out (something); to hand (something) out
(to do, to complete) homework
illustration(s); to illustrate
(to give) instruction(s); to instruct
(academic, personal) journal(s)
keyboard(s); to key/type in (data), to key/type (data) in
knowledge; to know; (to do something) knowingly
(to give, to attend, to listen to) a lecture (lectures); to lecture
login(s); to log in
logout(s); to log out
mark(s); to mark (a paper)
(magic, whiteboard) marker(s)
to miss (a class, a quiz); (to be) missing

(to make) a mistake (mistakes);
to mistake (one thing or person) for (another thing or person); mistaken; mistakenly
(to express, to give, to offer) an opinion (opinions)
(to write, to do, to give, to present) a paper (papers)
paragraph(s); to paragraph
(class) period(s); periodic(al) (test); periodically (review)
plagiarism; to plagiarize
point(s) (on an exam, in a course)
practice; to practice
printout(s); to print out (something), to print (something) out
(to have, to solve) a problem (problems); problematic
publication(s); to publish
(to ask) a question (questions); to question
(to give, to take) a quiz (quizzes); to quiz
to raise (one's) hand
(to do, to conduct) research; to research
(to give) a/some response; to respond
requirement(s); to require
review(s); to review
schedule(s); to schedule (something) on (date) for/at (time)
semester(s)
seminar(s)
sentence(s)
solution(s); to solve
study (studies); to study; studious; studiously
syllabus (syllabi/syllabuses)
(to take, to have, to give) a test (tests); to test
text(book)(s)
theory (theories); to theorize; theoretical; theoretically
(sub)title(s); to (sub)title
university (universities)
whiteboard(s)
(to do) work; to work
workshop(s)
workstation(s)

3.2 University of Aizu Physical Environment Vocabulary

administration complex/building
athletic field
auditorium
bus stop
classroom(s)
clubroom(s)
computer assisted instruction room(s), CAI room(s)
computer lab(s)
conference room
(quadrangle) courtyard(s)
drinking/water fountain
energy center

field house
gymnasium
language media lab(oratory)/lab(oratories)s, LML(s)
lecture hall
library lounge(s), research library/libraries
martial arts hall
men's/women's locker room(s)
(main) library
monitor(s)
mouse (mice)
multimedia center
office(s)
parking lot(s)
printer(s)
relaxation room (= SLRU 1)
research lab(s)
research quadrangles
men's/women's restroom(s)
(computer, TV) screen(s)
stairwell(s)
student hall
tennis court(s)
track
university bookstore
university cafeteria
university park
university pool
university restaurant
university shop
waterside park
weightroom (= SLRU 2)

3.3 University of Aizu People Vocabulary

Common Noun Forms	Proper Noun Forms (Official Titles)
university president	President, University of Aizu
director of the department for student affairs	Director, Department for Student Affairs
department head(s)	Director, Department of Computer Hardware
	Director, Department of Computer Software
center head(s)	Director, Center for Language Research
	Director, Center for Mathematical Sciences
	Director, Center for Cultural Research and Studies
(university) faculty	
= professors, associate professors, assistant professors, and research associates	
professors	
= professors, associate professors, assistant professors	
(full) professor(s)	Professor(s)
associate professor(s)	Associate Professor(s)
assistant professor(s)	Assistant Professor(s)
research associate(s)	Research Associate(s)
office staff	
secretary (secretaries)	
food staff	
sales clerks/staff	
maintenance personnel	
security guard(s)	
graduate student(s)	
undergrad(uate) (student)(s)	
senior(s)	
junior(s)	
sophomore(s)	
freshman/freshmen	

3.4 University of Aizu Computer Science Discourse Vocabulary

The following lists of noun, gerund, verb, and adjective forms were constructed after an analysis of lexical frequency was run on the *University of Aizu 1993 Annual Review: School of Computer Science and Engineering*. Technical items that primarily appear in professional computer science discourse (and thus requiring definition in computer science glossaries) are marked with a double asterisk (**). Semi-technical items that the computer science community has adopted from general English and added special meanings to are marked with an asterisk (*). All other items are of a non-technical nature (i.e. general English) but used often in computer science discourse.[3] Other parts of speech, such as adverbs, articles, and prepositions, have not been listed since the items that appeared under these categories are already known by most incoming freshmen.

[3]Vocabulary that has been listed in other categories, such as the list of University of Aizu Academic Vocabulary, have not been repeated here.

3.4.1 Nouns and Gerunds

computer(s), system(s), software, design(s), model(s), method(s), *processing, achievement(s), summary (summaries), data/datum (data), *algorithm(s), information, recognition, project(s), development(s), *modeling, *language(s), time, result(s), performance(s), conference(s), application(s), *hardware, *network(s), science(s), simulation(s), approach(es), management, society (societies), use(s), communication, *computing, circuit(s), control(s), device(s), *multimedia, case(s), *architecture, engineering, graph(s), process(es), specification(s), **VLSI (Very Large Scale Integration), *graphics, shape(s), application(s), audio, number(s), work, state(s), technology (technologies), image(s), program(s), structure(s), technique(s), current, electron(s), object(s), physics, sound, speech, group(s), space(s), member(s), environment(s), *interface(s), *programming, set(s), surface(s), order(s), property (properties), circuit(s), logic, memory, *processor(s), support, activity (activities), behavior, function(s), **CAD (= Computer-Aided Design), point(s), r presentation(s), world(s), agent(s), courseware, learning, nuclei, base(s), *database(s), hand(s), machine(s), reality, semantics, synthesis, computation(s), feature(s), function(s), implementation, mathematics, *protocol(s), *editor(s), form(s), goal(s), action(s), character(s), evaluation(s), verification, visualization, effect(s), generation, term(s), tool(s), animation, area(s), basis (bases), complexity (complexities), concept(s), description(s), search(es), *semigroup(s), signal(s), *tree(s) (= diagram), component(s), course(s) (= direction), element(s), foundation(s), optimization(s), **semiconductor(s), user(s)

Note: These nouns (the singular form, the plural form, or both forms) or gerunds appeared 20 or more times in the text. They are listed in descending order of frequency.

3.4.2 Verb Forms

be, design, model, has, base, process, use, develop, distribute, test, study, compute, control, propose, graph, work, state, *program, structure, give, order, show, support, consider, proceed, learn, apply, follow, present, include, obtain, transport, form, route, call, investigate, patch, search, signal, relate, advance, allow, gain, make, provide, accept, implement, press, report, decentralize, speed up, know, need, train, change, cross, define, evaluate, exist, delay, invite, suggested, understand, balance, connect, describe, feature, planning, point, review, time, window, account for, branch, find, like, open, transfer, free, generate, map, practice, queue, service, specify, teach, view, write, cluster, flow, mean, operate, perform, propose, truncate

Note: These verbs (or various forms of these verbs) appeared 10 or more times in the text. They are listed in descending order of frequency.

3.4.3 Adjective Forms

parallel, new, international, such, mathematical, human, some, other, complex, each, self-timed, asynchronous, all, current, more, both, various, any, first, virtual, high, scientific, several, general, computational, many, technical, academic, neural, basic, geometric, important, main, continuous, different, educational, structural, efficient, global, spatial, fast, formal, geometrical, large, **object-oriented, real, visual, quantum, dynamic, possible, simple, infrared, joint, theoretical, second, disjoint, finite, single, common, far, optimal, particular, physical, abstract, better, algebraic, effective, local, multiple, synthetic, artificial, conventional, cultural, experimental, genetic, handwritten, light, natural, necessary, potential, singular, bipolar, *fuzzy, graphical, major, nuclear, adaptive, logical, national, open, standard, topological, concurrent,

free, good, intelligent, interactive, numerical, arbitrary, autonomous, available, certain, dimensional, hierarchical, hot (= popular), minimum, next, nonlinear, powerful, same, specific, wide

Note: These adjectives (or noun forms functioning primarily as adjectives) appeared 10 or more times in the text. They are listed in descending order of frequency.

3.5 Miscellaneous Research Statistics

lexical base:[4] 51,983 items

approximate portion of lexical base already known[5] by students: 4,109 items (7.9 %)

lexical range:[6] 7,683

approximate percent of lexical range already known by students: 624 items (8 %)

approximate percent of lexical base to be taught: 11,200 items (21.5 %)

approximate percent of lexical range to be taught: 640 items (8.3 %)

approximate number of additional items to be taught: 330 items

approximate number of total items to be taught: 970 items[7]

percentage of 970 items that are technical: 3 (0.3 %)

percentage of 970 items that are semi-technical: 30 (3 %)

percentage of 970 items that are non-technical (general English): 937 (96.7%)

This study not only identified items that are appropriate for overt instruction in freshman Composition 1 and 2 (the course primarily responsible for attention to vocabulary, grammar, and introductory writing instruction), but this project also confirmed our suspicions that the majority of words that students need to know to negotiate their way through computer science content material are words common to general English rather than words of a semi- or highly technical nature. The following statistics illustrate this point:

lexical base: 51,983 items

lexical range: 7,683 different items

percent of lexical base that is technical: approx. 200-300 (0.4-0.6%)

percent of lexical range that is technical: approx. 180-250 (2.3-3.3%)

percent of lexical base that is semi-technical: 1550-1650 (3.0-3.2%)

percent of lexical range that is semi-technical: 900-1200 (11.7-15.6%)

percent of lexical base that is non-technical: 50,000-50,200 (96.2-96.6%)

percent of lexical range that is non-technical: 6200-6600 (81.1-86.0%)

[4] The lexical base is composed of all common and proper nouns, initials, abbreviations, and acronyms in the *Annual Review*.

[5] We assume that all words normally taught in Japanese junior and senior high schools constitute words *already known* by students.

[6] The lexical range is the number of *different* lexical items.

[7] Singular forms, plural forms, and changes in the verb form each constitute a separate item.

For introductory language instruction at the freshman level, it is obvious that concentration on the general English vocabulary most common to academic instruction and computer science discourse would prioritize English language instruction in the most logical and efficient way.

4 Educational Applications

Though the intent of this report is to identify and report the English vocabulary that is necessary for incorporation into freshman English composition instruction, it would be appropriate here to suggest briefly how this corpora might best be taught.

Clearly, vocabulary is learned through association, both contextual and relational. Research has demonstrated that human beings "learn" a word by associating the written and/or spoken forms with the context in which the word is "experienced" and/or by associating it with other items in the learner's personal lexicon (Hayes-Roth, B. & Hayes-Roth, F., 1977; Lehrer, A. 1978; Meara, P., 1983, 1984; Carter & McCarthy, 1988; Gass & Selinker, 1994). If we want vocabulary instruction in Composition 1 and 2 to be successful, we must be sensitive to this principle. This means that we must *use* the vocabulary in natural contexts, both in the dissemination of content material about academic life and computer science and in course assignments that encourage students to employ these words to convey their thoughts to one other in class. We envision a combination of problem-solving exercises, student-teacher journaling, essay writing, and electronic correspondence that will provide plentiful opportunities for reading and using these words in natural communicative contexts. In addition, we plan to experiment with various vocabulary ideas presented in one of TESOL's latest publications, *New Ways in Teaching Vocabulary* (Nation, 1994) to see which activities are applicable to our situation here at the University of Aizu. Also, pre- and post-tests will be employed to chart progress and monitor the success of both students and the instructional activities. Then finally, in a year from now, a follow-up report will document our findings and suggest revisions to the curriculum for the following year. If such a cyclical process can be continued year-after-year, only then can thoroughly informed and effective composition instruction be guaranteed.

Acknowledgement: Professors Thomas Orr, Kiel Christianson, and Christian Goetze would like to express a note of thanks to Hideaki Okawara for assisting us in many small details of this project. One of our goals at the University of Aizu is to produce university research professors, an occupation Mr. Okawara personally seeks. It is our hope that this on-the-job training in one small bit of university research and publication will not only begin to build his personal vita but also begin to lay the foundation of experience that may prove useful for future activity in research and publishing. Hideaki, on this day, your 20th birthday, we wish you success in your pursuit of knowledge and expertise in computer science.

References

Bramki, D. & Williams, R. (1984). Lexical familiarization in economics texts and its pedagogical implications in reading comprehension. *Reading in a Foreign Language 2*, 169-81.

Chandler-Burns, R. (1995). English corpora for science and technology. In T. Orr, (Ed.) *English for science and technology: Profiles and perspectives* (pp. 9-15). Aizuwakamatsu, Japan: University of Aizu.

Chandler-Burns, R. et al. (1994). Spanish 92: the first computational linguistic analysis of contemporary Spanish for medical purposes. Available via ftp @ xxx.lanl.gov/cm-lg/9404006 and ftp @ titan.fullerton. edu/pubs/research/chandler/ascii/S92 and E92.

Chandler-Burns, R et al. (1994) Un nuevo paradigma para pronosticar funcion a partir de estructuras primarias de genes y proteinas: hacia una gramatica biolinguistica (primera comunicacion). Monterrey: *Memorias del XII Encuentro de Investigacion Biomedica*, Publicaciones de Subdireccion de Postgrado, p. 121.

Chandler-Burns, R et al. (1987). Computational linguistic analysis of the medical lexis. *EMP Newsletter*, 4(2) 44-47.

Chandler-Burns, R. et al. (1986). *Analisis linguistico computacional del idioma ingles actual para propositos medicos: el corpus universidad autonoma de Nuevo Leon (CUANL).* Monterrey: Publicaciones de Subdireccion de Postgrado/3er lugar Premio de Investigacion 1986.

Carter, R. & McCarthy, M. (1988). *Vocabulary and language teaching.* London: Longman.

Flowerdew, J. (1993). Content-based language instruction in a tertiary setting. *English for Specific Purposes, 12*, 121-138.

Gass, S. & Selinker, L. (1994). *Second language acquisition: An introductory course.* Hillsdale, NJ: Lawrence Erlbaum. Hayes-Roth, B. & Hayes-Roth, F. (1977). The prominence of lexical information in memory representations of meaning. *Journal of Verbal Learning and Verbal Behavior 16*, 119-36.

Inman, M. (1978). Lexical analysis of scientific and technical prose. In M. Trimble, L. Trimble, & K. Drobnic (Eds.) *English for specific purposes: Science and technology* (pp. 242-56). Corvallis, Oregon State University.

Kucera, H. & Francis, N. (1967). *Computational analysis of present-day American English.* Providence, Rhode Island: Brown University Press.

Lehrer, A. (1978). Structures of the lexicon, and transfer of meaning. *Lingua 45*, 95-123.

Meara, P. (1984). The study of lexis. In A. Davies, C. Criger, and A. Howatt (Eds). *Interlanguage* (pp. 225-35). Edinburgh University Press.

Meara, P. (Ed.). (1983). *Vocabulary in second language.* London: CILT.

Nation, P. (1994). *New ways in teaching vocabulary.* Alexandria, VA: Teachers of English to Speakers of Other Languages.

Public Relations and Publication Committee. (1994). *University of Aizu 1993 annual review: School of computer science and engineering.* Aizuwakamatsu, Japan: University of Aizu.

Salager-Meyer, F. (1983). Fundamental medical vs. general lexis: A frequency analysis. *Reading in a Foreign Language*, 1(1) and *EMPN*, 4(1) 58-60. Ph.D. dissertation University of Texas at Austin.